The final publication is available through IOP Publishing Ltd. via http://dx.doi.org/doi:10.1088/1361-6501/aa5cfd

Rapid, high-resolution measurement of leaf area and leaf orientation using terrestrial LiDAR scanning data

Brian N. Bailey

University of California, Davis, Department of Plant Sciences, One Shields Ave., Davis, CA, 95616 USA

E-mail: bnbailey@ucdavis.edu

Walter F. Mahaffee

USDA-ARS, 3420 NW Orchard Ave., Corvallis, OR 97330 USA

E-mail: walt.mahaffee@ars.usda.gov

Abstract.

The rapid evolution of high performance computing technology has allowed for the development of extremely detailed models of the urban and natural environment. Although models can now represent sub-meter-scale variability in environmental geometry, model users are often unable to specify the geometry of real domains at this scale given available measurements. An emerging technology in this field has been the use of terrestrial LiDAR scanning data to rapidly measure the three-dimensional geometry of trees such as the distribution of leaf area. However, current LiDAR methods suffer from the limitation that they require detailed knowledge of leaf orientation in order to translate projected leaf area into actual leaf area. Common methods for measuring leaf orientation are often tedious or inaccurate, which places constraints on the LiDAR measurement technique. This work presents a new method to simultaneously measure leaf orientation and leaf area within an arbitrarily defined volume using terrestrial LiDAR data. The novelty of the method lies in the direct measurement of the fraction of projected leaf area G from the LiDAR data which is required to relate projected leaf area to total leaf area, and in the new way in which radiation transfer theory was used to calculate leaf area from the LiDAR data. The method was validated by comparing LiDAR-measured leaf area to 1) 'synthetic' or computer-generated LiDAR data where the exact area was known, and 2) direct measurements of leaf area in the field using destructive sampling. Overall, agreement between the LiDAR and reference measurements was very good, showing a normalized root-mean-squared-error of about 15% for the synthetic tests, and 13% in the field.

Keywords: LiDAR, Leaf area, Leaf angle distribution function

Table 1: List of primary symbols used throughout the text

Symbol	Description	Units
A	One-sided surface area	m^2
a_L	One-sided leaf area per unit volume (i.e., leaf area density)	m^{-1}
d	Index of agreement (Willmott, 1981)	-
G	Fraction of leaf area projected in a given direciton	-
$\mathcal{N}$	Total number of scan rays that traversed a voxel without intersecting a leaf	-
$\vec{n}_L$	Unit vector normal to the leaf surface	-
nRMSE	Normalized root-mean-squared-error	-
N_r	Total number of scan rays that traversed a voxel	-
N_t	Total number of triangles in a voxel	-
P	Probability that a ray is not intercepted	-
r	Length of a segment along ray path	m
$\vec{r}$	Unit vector in the direction of radiation propagation	
t	Parametric distance along ray path relative to the scan origin	m
κ	Attenuation coefficient	m^{-1}
θ	Ray zenith angle	rad
φ	Ray azimuth angle	rad

1. Introduction

With the continued improvement of high-performance computing technology, high-resolution models are becoming an increasingly prevalent means for understanding complex, multi-scale processes in the environment. Applications of such models include

weather prediction (Skamarock et al., 2005; Kerr, 2012), understanding greenhouse gas exchanges and their contributions to the global climate (Searchinger et al., 2008), and urban design and planning (Fong et al., 2009; Wu et al., 2010). These environmental models have seen radical improvements in their ability to represent fine-scale details associated with various transport processes of interest. However, increasingly detailed models inevitably require increasingly detailed measurements of model inputs. Among the most important of these inputs is the accurate measurement and specification of object geometry.

Within environmental models (e.g., Lindberg and Grimmond, 2011; Lemonsu et al., 2012; Krayenhoff et al., 2014; Bailey and Stoll, 2016), plants often create heterogeneous geometry that is difficult to specify. Since it is usually not feasible to represent individual leaves in models, plant leaves are commonly represented statistically by specifying an effective volume of leaves with a given leaf area density (LAD) and probability density function (p.d.f.) of leaf orientation (Ross, 1981; Bailey et al., 2014).

Recent environmental transport models are able to represent variability in LAD and the leaf orientation p.d.f. over meter-scales or smaller (e.g., Sinoquet et al., 2001; Olchev et al., 2009; Kobayashi et al., 2012; Bailey et al., 2014, 2016). However, obtaining measurements of these model inputs at such fine scales is a formidable challenge. Manual measurement is the most straightforward method for measuring the LAD and leaf orientation p.d.f. Leaf area can be measured directly by removing all leaves in a designated volume and summing their surface area (Jonckheere et al., 2004). The leaf orientation p.d.f. can be populated using a compass and protractor to measure the orientation of many leaves (Ross, 1981; Norman and Campbell, 1989). Finally, the overall size and shape of a given plant can be directly measured, or inferred based on its shadow.

While these direct methods provide a straightforward means of obtaining geometric inputs for plant-related environmental models, they are typically far too time consuming to provide data that describes meter-scale variability (Norman and Campbell, 1989), which modern models can effectively utilize. To provide more efficient measurements of geometric inputs, workers have attempted to use indirect methods that more rapidly infer geometry. These methods include but are not limited to gap fraction analysis (Norman and Campbell, 1989; Welles and Norman, 1991; Bréda, 2003), inclined point quadrats (Warren Wilson, 1960; Caldwell et al., 1983), and spectral analysis of remotely sensed data (Olchev et al., 2009; Jones and Vaughan, 2010). While these methods are typically much more efficient and less destructive than manual measurements, they all make compromises between accuracy, effort, and level of spatial description.

Terrestrial or ground-based LiDAR scanning is a relatively new measurement technology that shows much potential for rapidly providing sub-meter-scale measurements of plant geometry. The LiDAR instrument is able to accurately measure the distance to millions of points surrounding the scanner location, which has been successfully used to measure leaf area contained within an arbitrary volume (Henning and Radtke, 2006; Hosoi and Omasa, 2006; Rosell et al., 2009; Yang et al., 2013; Béland et al., 2014). By

subdividing a plant into many sub-volumes, the method can resolve meter-scale variability in leaf area. As a byproduct, the method also specifies the size and shape of the plant, since volumes outside of the plant will have a leaf area of zero (Béland et al., 2014).

The accuracy and utility of terrestrial LiDAR measurement of leaf area has been limited by several necessary assumptions. First, is that the sub-volumes of leaves for which LAD is calculated must be sufficiently small that the amount of occluded or hidden leaf area within a volume is minimal (Béland et al., 2014). Second, in order to infer LAD, the leaf orientation distribution p.d.f. must be known, which means that it must either be measured independently or a simplifying assumption must be made. Performing such measurements has traditionally involved manual sampling of leaf orientations to generate the p.d.f., which can be time consuming. To avoid the need for collecting such measurements, it is common to assume that leaves are randomly oriented, but such a leaf angle distribution is rarely present in nature (Pisek et al., 2013). Bailey and Mahaffee (2017) introduced a method that can directly estimate the leaf orientation p.d.f. from LiDAR scanning, which has the potential to improve LiDAR leaf area measurements.

The present work overcomes the above limitations associated with LiDAR leaf area measurement by 1) developing new theory for LAD inversion consistent with radiation transfer theory that reduces required theoretical assumptions, and 2) rapidly measuring the leaf orientation p.d.f. directly from the LiDAR scan and coupling it with the LAD inversion procedure. The resulting method provides a means for rapidly and simultaneously measuring the leaf area and leaf orientation p.d.f. for an arbitrary volume of leaves.

2. Materials and Methods

2.1. Terrestrial LiDAR scanning

Terrestrial or ground-based LiDAR scanners emit a very large number of concentrated radiation pulses in the spherical space surrounding the scanner. In the event that the pulses intersect an object, some fraction of the radiation is scattered back to the scanner which is detected by an on-board sensor. LiDAR scanners use various techniques to determine the distance to the object that was hit by the pulse, including 'time-of-flight', which uses the time that it takes for the pulse to scatter and return to the sensor to determine the distance. Other scanners analyze the spectral signature of the scattered radiation to infer distance. Typical scans generate millions of ray-vegetation intersection points in a timeframe of several minutes that give the three-dimensional Cartesian coordinate to objects surrounding the scanner.

2.2. Measurement methodology overview

The overall strategy used in this work to measure leaf area using LiDAR scans was to invert the Beer-Lambert law for the attenuation coefficient, which can be related to the leaf area density. The Beer-Lambert law for the attenuation of radiation through a participating medium can be written as (Modest, 2013)

$$P = \exp\left(-\kappa r\right), \tag{1}$$

where P is the probability that a beam of radiation can travel a distance r without being intercepted, and κ is the attenuation coefficient of the medium. For a medium of leaves, κ can be interpreted as the one-sided surface area of leaves projected in the direction of radiation propagation per unit volume (Ross, 1981), or

$$\kappa = a_L G, \tag{2}$$

where a_L is the one-sided leaf area per unit volume (i.e., leaf area density, LAD), and G is the fraction of leaf area projected in the direction of radiation propagation (i.e., the so-called G-function; Ross, 1981). Note that LAD can be converted to total one-sided leaf area within a voxel by multiplying a_L by the voxel volume. The G-function ranges from 0 in the case of all leaves oriented parallel to the direction of propagation to 1 when all leaves are projected in the direction of propagation. If leaves have no preferred direction (randomly oriented), $G = 0.5$ regardless of the direction of propagation, meaning that exactly half of the total leaf area is projected in any direction. An assumption of $G = 0.5$ is commonly used in the literature for simplicity, but in reality G is a strong function of direction and location within plants (Pisek et al., 2013; Bailey and Mahaffee, 2017), and thus such an assumption was not made in this work.

Our goal is to apply Eq. 1 to an arbitrary volume or 'voxel' of leaves that has been interrogated by a LiDAR scan (Fig. 1), and invert this equation to give a_L. In order to do so, we must develop a method to apply Eq. 1, which is only valid for a single ray on average, to that of many rays traversing a given voxel with LAD of a_L. Additionally, we also must properly calculate P, G, and r in a manner consistent with the method.

2.3. Calculation of the G-function

Application of Eq. 1 along a ray path has the potential to yield κ, the *projected* leaf area density in the direction of ray propagation $\vec{r}$ (Fig. 2). In order to separate the *total* leaf area from κ, the fraction of total leaf area projected in the direction of $\vec{r}$ (denoted as G or G-function) must be calculated.

Traditionally the average G-function for a volume of leaves is calculated by first forming a p.d.f. of leaf orientation, and integrating the product of the p.d.f. and the dot product of the ray direction and the leaf normal ($\vec{r}\cdot\vec{n}_L$) over all possible leaf normals (see Ross, 1981). Below, we present a method that bypasses the need to explicitly calculate the leaf angle p.d.f., and instead calculates the average G-function directly from the LiDAR data.

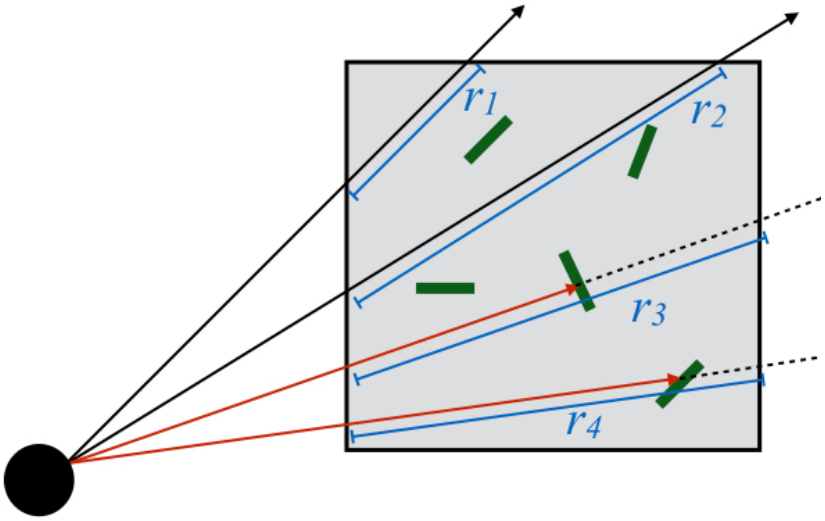

Figure 1. Schematic sketch depicting a simplified LiDAR scan in two dimensions. Rays originate from the scanner (filled black circle), which intersect a voxel containing randomly oriented leaves (green bars). Beams that intersect a leaf within the volume are coloured red, while beams that pass through the volume without intersecting any leaves are coloured black. The *total* intersection length between the voxel and a ray of infinite length is denoted by r_i. For the example illustrated, the probability of no ray intersection is $\overline{P} = \frac{2}{4} = 0.5$.

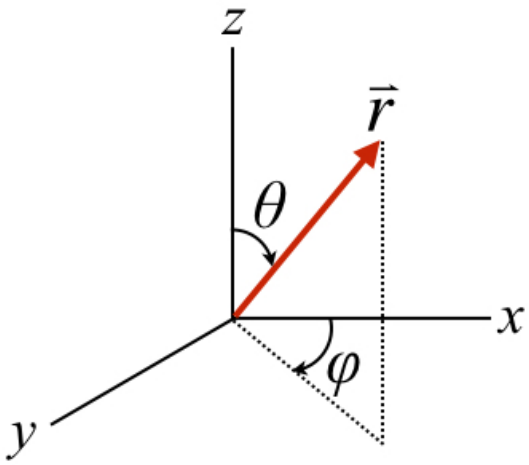

Figure 2. Schematic of the spherical coordinate system used.

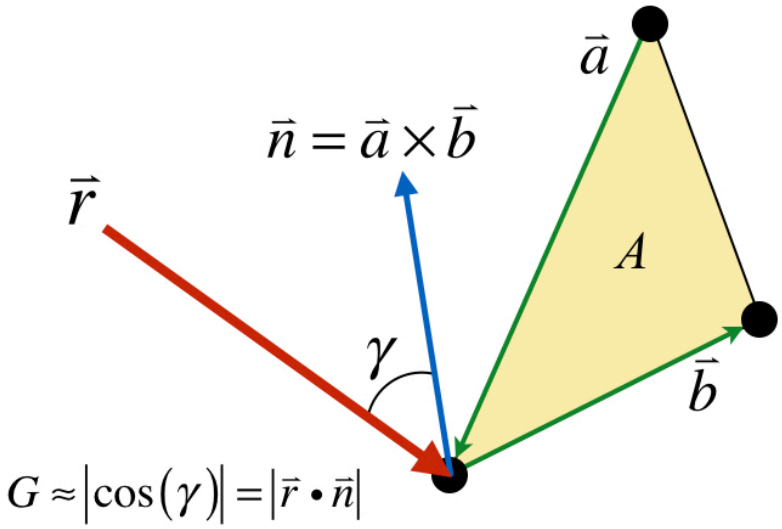

Figure 3. Schematic sketch depicting the calculation of the projected area of a triangle.

Leaf orientations can be measured using LiDAR data and the triangulation algorithm developed by (Bailey and Mahaffee, 2017). To summarize, adjacent ray-leaf intersection points are connected to form triangles, from which leaf normals $\vec{n}_L$ can be easily calculated (Fig. 3). The algorithm produces continuous triangles that approximately reconstruct leaf surfaces. Each triangle is associated with three ray-leaf intersection points that make up its vertices as shown in Fig. 3.

These triangles are used along with the following approach to calculate G. A unit vector pointing in the direction of each of the three rays $\vec{r}$ making up the vertices of a triangle is given by subtracting the (x, y, z) coordinate of the scan origin (i.e., location of scanner) from the triangle vertex (and normalizing by the magnitude). Although $\vec{r}$ varies slightly for each triangle vertex, it is assumed that $\vec{r}$ is equal across the whole triangle given that the solid angle subtended by the triangle with respect to the scanner is extremely small. Thus, each triangle is associated with one ray direction $\vec{r}$. The approximate fraction of the i^{th} triangle area A projected in the i^{th} ray direction $\vec{r}_i$ is given by

$$G_i = \left|\vec{r}_i \cdot \vec{n}_{L,i}\right|. \tag{3}$$

The average G-function $\overline{G}$ for a voxel can be calculated by averaging G_i over all triangles within the voxel. However, it is critical that the proper weighting be applied to each G_i value in the average. Firstly, G_i should be weighted by the triangle area to avoid biasing toward triangles with normals $\vec{n}_{L,i}$ closer to parallel with $\vec{r}_i$ (Bailey and Mahaffee, 2017). Secondly, G_i should be weighted by $\sin\theta_i$ to avoid biasing toward ray directions closer to vertical (Bailey and Mahaffee, 2017), where θ_i is the zenith angle of the ray direction (i.e., $\cos^{-1} r_z$, where r_z is the vertical component of $\vec{r}$). With these considerations in

mind, the average G-function for a volume is given by

$$\overline{G} = N_t \left\{ \frac{\sum_{i=1}^{N_t} G_i \, A_i \sin\theta_i}{\left(\sum_{i=1}^{N_t} A_i\right)\left(\sum_{i=1}^{N_t} \sin\theta_i\right)} \right\}, \tag{4}$$

where N_t is the total number of triangles in the voxel.

2.4. Calculation intersection probability $\overline{P}$

The remaining terms in Eq. 1 to be calculated are the probability that a ray is not intercepted in the volume P, and the volume traversal distance r. From a given ray it can be directly determined whether a leaf-ray intersection occurs after some distance r, which gives a binary value. To form a probability of intersection, the ensemble of all rays traversing a given voxel can be considered. The average probability that a ray will not intersect a leaf within a voxel can be estimated by the ratio of the number of rays traversing the voxel with no intersection (Fig. 1, black rays), to the total number of rays traversing the voxel (Fig. 1, all rays). Because of the scanning pattern of the LiDAR, one must avoid biasing toward rays with small zenith angles by weighting each ray by $\sin\theta_i$. Thus, the probability of a ray traversing a voxel without intersecting a leaf, $\overline{P}$, is given by

$$\overline{P} = \frac{\sum_{j=1}^{\mathcal{N}} \sin\theta_j}{\sum_{k=1}^{N_r} \sin\theta_k}, \tag{5}$$

where the sequence $j = 1, \mathcal{N}$ corresponds to rays that traversed the voxel without intersecting a leaf, the sequence $k = 1, N_r$ corresponds to all rays traversing the voxel.

To determine whether a given ray intersects a voxel, methods commonly used in ray-tracing can be used. Suffern (2007) provides an excellent description of ray-voxel intersection testing for axis-aligned voxels, along with example C++ code. For non-axis aligned voxels, an inverse transformation can be applied to the ray, which is also detailed in Suffern (2007).

The ray-voxel intersection tests and calculation of $\overline{P}$ proceeds as follows:

(i) For a given LiDAR ray direction $\vec{r}_i$, calculate the parametric distance t along the ray path between the scanner and the ray intersection point (if the ray intersected the sky, $t \to \infty$).

(ii) Intersect the ray direction and all leaf voxels. If the ray intersects a voxel, calculate the parametric t-distance to the point where the ray enters the voxel (t_0) and the point where the ray exits the voxel (t_1).

- If $t < t_0$, the ray intersected a leaf before entering the voxel (ignore the intersection).

- If $t > t_1$, the ray passed through the voxel without intersecting a leaf. Calculate $\sin\theta_i$ (for weighting) and the intersection distance $t_1 - t_0 = r$ (to be used in Sect. 2.5).
- If $t \geq t_0$ and $t \leq t_1$, the ray intersection occurred with in the voxel. Calculate $\sin\theta_i$ (for weighting) and the intersection distance $t_1 - t_0 = r$ (to be used in Sect. 2.5).

(iii) Given the previous intersection tests, Eq. 5 can be used to calculate $\overline{P}$.

For many voxels and a high number of scan points (e.g., millions), these ray-voxel intersection tests can become quite computationally expensive. To accelerate computations, we performed ray-voxel intersection tests on the machine's graphics processing unit (GPU). The algorithms were implemented using NVIDIA's compute unified device architecture (CUDA) framework. The intersection tests are what may be termed 'embarrassingly parallel', meaning that rays are completely independent. In our implementation, each ray corresponded to a single CUDA thread, which performed intersection tests for each voxel. In the event of an intersection, $\sin\theta_i$ was accumulated in one of two buffers (depending on whether the intersection was inside or beyond the voxel) using an atomic addition operation. Using this approach, ray-voxel intersection tests for ∼100 million rays (a typical dataset size, see below) took on the order of a few seconds to run.

2.5. Inversion to calculate leaf area

2.5.1. Point quadrat analogy method: The typical method for calculating a_L based on estimations of $\overline{P}$, $\overline{G}$, and $\overline{r}$ is to use an analogy to inclined point quadrat sampling (Radtke and Bolstad, 2001; Hosoi and Omasa, 2006; Béland et al., 2011). Using this analogy, one can write

$$\overline{P} = 1 - a_L \overline{r} \overline{G}, \tag{6}$$

which can be solved algebraically for a_L

$$a_L = \frac{1 - \overline{P}}{\overline{r}\overline{G}}. \tag{7}$$

It is known that this approach is limited by the assumption that there are no occluded leaves within a given voxel, or that a probe extending through the voxel is most likely to only intersect one leaf (Béland et al., 2014). This is because a ray can only intersect a single leaf, whereas a true point quadrat can continue on and intersect multiple leaves. Thus, most previous LiDAR measurements of leaf area have been limited to relatively small voxels (e.g., Radtke and Bolstad, 2001; Hosoi and Omasa, 2006; Béland et al., 2011, 2014).

2.5.2. Beer's law inversion method: As noted previously, the use of Eq. 1 to calculate a_L overcomes many of the limitations associated with the 'point quadrat' approach described in Sect. 2.5.1. Although for a single ray Eq. 1 represents a scalar equation, this equation applied to an ensemble of rays intersecting a volume is a vector equation. Some type of aggregation or averaging must be applied to allow for inversion. The simplest of which is the following

$$\overline{P} = \exp\left(-a_L \overline{G}\, \overline{r}\right), \tag{8}$$

where, here $\overline{r}$ is the average ray-voxel intersection distance for all rays passing through a given volume. This equation can be solved algebraically for a_L

$$a_L = \frac{-\ln \overline{P}}{\overline{r}\overline{G}}. \tag{9}$$

Note that even for rays that intersected leaves within a voxel, the traversal distance r should be the total distance from voxel entrance to exit as calculated in the previous section (i.e., $r = t_1 - t_0$). This is because Eq. 8 should be interpreted as follows: $\overline{P}$ is the probability of a ray traversing a distance of $\overline{r}$ through a projected leaf area density of $a_L\overline{G}$ without intersecting a leaf.

If $\overline{P}$ is near 1, such as when the voxel size or leaf area is small, Eq. 9 is nearly the same as Eq. 7. Equation 7 should always give a lower value of a_L than Eq. 9 since 1-$\overline{P}$ is always less than $-\ln \overline{P}$ given that $0 \leq \overline{P} \leq 1$.

Although Eq. 8 is a practical way of aggregating all rays intersecting a voxel, the averaging methodology presents problems if r varies substantially over a given voxel. In this case, r should be exponentially weighted in terms of the average, i.e.,

$$\overline{P} = \overline{\exp\left(-a_L \overline{G}\, r\right)} = \frac{1}{N_r} \sum_{k=1}^{N_r} \exp\left(-a_L \overline{G} r_k\right), \tag{10}$$

This difficulty with this equation is that an explicit solution for a_L is not possible, and thus an iterative solution is required. Any number of standard iterative equation solvers could be used for this task (e.g., the secant method, which was used in our implementation; Press et al., 2007), each of which essentially involves substituting 'guesses' for a_L into Eq. 10 until the equation balances.

It is expected that Eq. 8 will always give a smaller value of a_L than Eq. 10, since the average of $\overline{r}$ in Eq. 8 biases toward larger values of r_i. Both averaging methodologies will be tested in the following section.

3. Validation and Discussion

3.1. Synthetic scan data

Validating LiDAR measurements in the field is a considerable challenge. Firstly, collecting manual measurements for comparison is very time consuming and requires destructive measurements, which limits the number of replicate experiments that can

be performed. Secondly, there are numerous sources of compounding errors, making it impossible to accurately separate errors in the LiDAR measurements and manual measurements. Thus, there is no 'gold standard' against which to compare. In an attempt to remedy these difficulties, we present a novel validation technique in which we use 'synthetic' or computer-generated LiDAR data that replicates actual LiDAR scans that would be performed in the field. The advantage of using synthetic data, is that we know the *exact* geometry being scanned, which means that all errors lie in the data processing algorithm. Clearly, the geometry is artificial and does not represent reality, but it is a useful first check before performing validation against field data (Sect. 3.2).

The synthetic scanning methodology involved launching simulated rays from a given point and intersecting them with virtual leaves of known orientation and area. Circular disks were chosen as a model for actual leaves because simple ray-object intersection tests could be performed (cf. Suffern, 2007). Rays were launched in the same pattern as is done by the LiDAR instrument described above, which were intersected with each circular disk to determine the closest hit point (or that no hit occurred). The disk area density was then calculated from the data in the manner described above.

A $1 \times 1 \times 1$ m^3 volume of 10 cm diameter disks was considered, which was 3 m away from the simulated scanner and situated on the ground (Fig. 4). The scanner location was 0.5 m above the ground. Disks were inserted into the volume with a uniform random position. The elevation and azimuthal angles of the disks were chosen by also drawing from a uniform distribution. A uniform distribution was chosen so that equal weights are given to all orientations to avoid favoring certain orientations. Cases were simulated that had 3^3, 4^3, 5^3, and 6^3 disks in the volume. Given the size of the volume ($1 \times 1 \times 1$ m^3) and the area of the disks ($\pi(0.05)^2$ m^2), the one-sided area density of the cases were 0.021, 0.50, 0.98, and 1.70 m^2/m^3. Since the random component introduces variation for each realization, 20 independent realizations were generated for each disk density. The scan density was chosen such that it was the same as the field scans described below in Sect. 3.2. As in Bailey and Mahaffee (2017), the triangulation parameter L_{max} was chosen to be 0.05 m throughout this paper.

3.1.1. G-function Box-and-whisker plots are shown in Fig. 5, which give the normalized error in the synthetic LiDAR measurement of $\overline{G}$ for the four densities considered. The average error in the prediction of $\overline{G}$ was relatively consistent for the range of densities considered, with the average error across all realizations and densities being about a 14% over prediction. It is expected that the method will slightly over predict $\overline{G}$, given that disks with normals nearly perpendicular to the ray path may not be fully triangulated. As the disk density increases, the variance between individual realizations tends to decrease, which is likely due to the fact that with a sufficient number of disks the random influence becomes less significant. When the number of disks is low, relatively rare cases when the triangulation fails may have a significant impact on $\overline{G}$. Or there may be cases in which we are 'lucky' and the triangulation is almost always successful. For a high number of disks, this random variation decreases, and triangulation failure

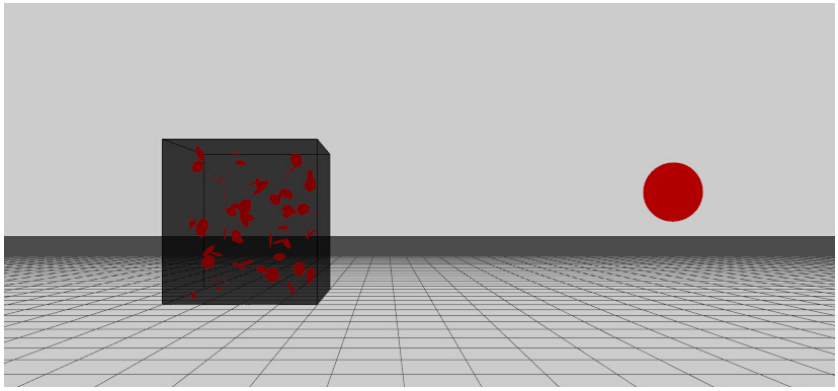

Figure 4. Visual depiction of the simulated scan with 64 circular disks. The red sphere denotes the scanner location, and the gray shaded cube denotes the interrogation volume. Disks are visualized using the hit point triangulation.

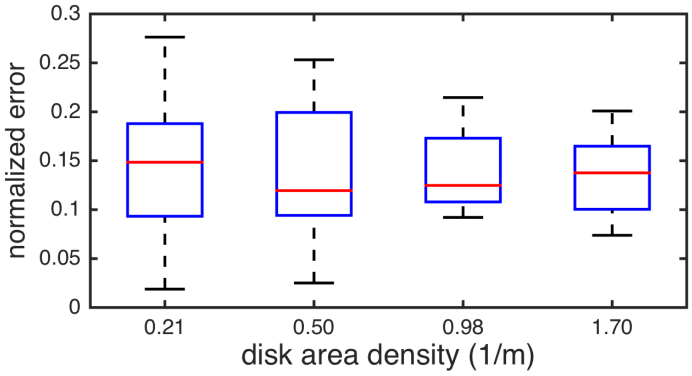

Figure 5. Box-and-whisker plots of normalized error in the predicted G-function for synthetic LiDAR data. The normalized error is defined as $(\overline{G}_{meas} - \overline{G}_{exact})/(\overline{G}_{exact})$, where $\overline{G}_{meas}$ and $\overline{G}_{exact}$ are respectively the measured and exact average G-functions. The red horizontal line denotes the median value of 20 realizations, the box denotes ± 1 standard deviation, and dashed brackets encompass the full range of values.

is more consistent and predictable.

3.1.2. Leaf Area Box-and-whisker plots are shown in Fig. 6, which give the normalized error in the synthetic LiDAR measurement of disk area. For all disk densities considered in this study, the radiative transfer theory approach (Eq. 10) consistently under predicted the disk density by roughly 15% on average. This under prediction error is roughly the same as the over prediction error in $\overline{G}$. An over prediction in $\overline{G}$ is likely to result in an under prediction in disk area, since the two are inversely related. For low disk area, $\overline{G}$ and disk area are approximately inversely proportional, and thus it is probable that most of the error in the disk area prediction is caused by error in the prediction of $\overline{G}$.

For both the point quadrat analogy (Eq. 6) and simple averaging scheme (Eq. 8), the normalized error continuously increased as the disk density was increased. For the

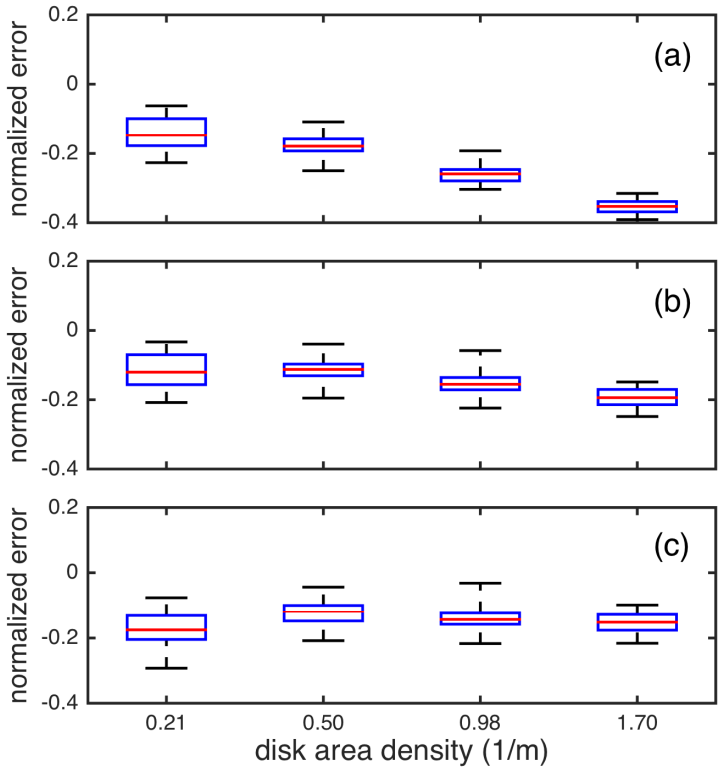

Figure 6. Box-and-whisker plots of normalized error in predicted leaf area for synthetic LiDAR data. The normalized error is defined as $(A_{meas} - A_{exact})/(A_{exact})$, where A_{meas} and A_{exact} are respectively the measured and exact disk areas. Different subplots correspond to different methodologies for calculating leaf area: (a) Eq. 6, (b) Eq. 8, and (c) Eq. 10. The red horizontal line denotes the median value of 20 realizations, the box denotes ±1 standard deviation, and dashed brackets encompass the full range of values.

point quadrat analogy, the negative bias was as large as 35% on average for the highest disk density. These tests confirm the expected result that Eqs. 6 and 8 tend to under predict leaf area compared to Eq. 10, an effect that becomes increasingly significant as disk (or leaf) area density increases.

3.2. Experimental data

To test the method for measuring leaf orientation and leaf area in the field, a validation experiment was conducted that focused on an isolated 6 m tall broad-leafed black cottonwood tree (*Poplus balsamifera*) in Oregon (44° 34' 11.11" N, 123° 14' 47.53" W). This tree was chosen because it represented the range of possible leaf inclination classes, it has relatively dense leaves (a challenge for LiDAR measurement), and because it was small enough to allow for the collection of direct measurements without the use of special equipment such as scaffolding.

In this study, the Focus^{3D} X 330 (FARO Technologies, Lake Mary, FL, USA) portable laser scanner was used to map the structure of vegetation. The scanner emits

laser pulses with a wavelength of 1550 nm in a uniformly gridded pattern over a field of view of 300° in the vertical and 360° in the horizontal (i.e., $\theta = 0$ to 150° and $\varphi = 0$ to 360°, see Fig. 2). The beam diameter at exit is roughly 2.25 mm and diverges at a typical half angle of about 0.011°. Upon intersection with an object (range from 0.6 m up to 330 m), the three-dimensional coordinates of the intersection are calculated and recorded, with a typical distance accuracy of ±2 mm. In the event that the beam intersects multiple objects (because of its non-negligibly small diameter), the scanner records only one distance measurement, which is effectively a weighted average of the distance to all objects that were hit. The scanner also collects a large number of color digital photographs during the scan. This allows a mapping of the 3D hit point onto the 2D photographs, which gives an RGB color value associated with each hit point.

In the experiment, vegetation was first scanned with the LiDAR at a density of 3415×8120 points (zenith×azimuth), which represents a moderate density for this instrument (this is 1/5 of the highest possible density). The scanner was placed approximately 2 m from the nearest vegetation, such that it was as close as possible to the tree while still having a direct line of sight to the majority of the upper crown and being within the minimum range of the instrument (0.6 m). At this scan density and range, adjacent hit points had a spacing that were approximately in the range of 1mm to 10 mm. The characteristic leaf dimension was on the order of 10 cm, with some leaves as large as 15-20 cm. A total of four scans were performed with the scanner at each of the four cardinal directions (see Fig. 7b), which took roughly 15 minutes each. To combine the four scans, they were first co-registered to a common coordinate system using the on-board GPS and inclinometer and using common objects in the scene, then the points were combined in the leaf area calculations by simply adding additional scans to the running sums in Eqs. 8 and 10.

Vegetation was divided into 12 discrete zones of size $2 \times 2 \times 1.1$ m^3, which were marked using high visibility tape. The tree was divided into three vertical levels, with four azimuthal zones at each level that faced the scan locations (Fig. 7). The first vertical level began at 0.76 m from the ground to avoid interference from grass and weeds. This left approximately 2 m of vegetation at the top of the tree that was not sampled because it could not be easily reached using available equipment. For future reference, volumes were numbered starting from North and moving clockwise, with voxels 1-4 being closest to the ground and voxels 9-12 furthest from the ground. By combining the four scans, this ensured that all volumes had a very large number of hit points, which avoided the need to establish a hit count threshold (Béland et al., 2014).

To reduce effort, destructive measurements were collected for leaf area in only 6 of the 12 zones (zones 1, 3, 6, 8, 9, 11), which were chosen in an alternating pattern. Leaves were stripped and bagged according to zone, then immediately transported back to the lab and refrigerated to minimize changes in shape/area. An LI-3100C area meter (LI-COR, Lincoln, NE, USA) was used to measure the total leaf area for each zone. All leaf area measurements were performed within a week of leaf stripping, which meant that changes in leaf area before measurement was minimal. The accuracy of the LI-3100C

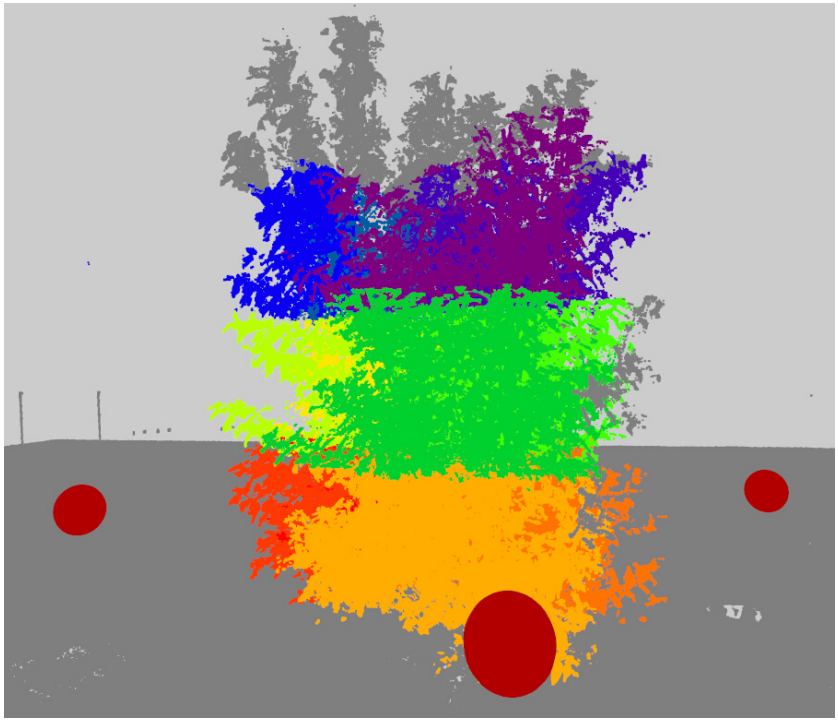

Figure 7. North-facing view of LiDAR scan point cloud. Tree is divided into 12 zones, and intersection points are coloured according to the voxel in which they reside (points falling outside of all voxels are coloured grey). Each of the scan locations are denoted by red spheres (Northern sphere is obstructed).

instrument is expected to be roughly ±5%. It is also important to note that the area of woody plant material was not included in the direct measurements, although the woody area is technically included in the LiDAR measurement. For the tree considered in this study, it seemed reasonable to neglect the contributions of woody material, although they could be included using, e.g., the method of Béland et al. (2014) if necessary.

Agreement between manual and LiDAR measurements was quantified using three statistical parameters: the index of agreement d, the normalized root-mean-squared-error (nRMSE), and mean bias. The index of agreement was used in lieu of more traditional indices (e.g., coefficient of determination R^2, Pearson correlation coefficient r) because it is expected to be a more robust indicator of agreement between two variables (Willmott, 1981, 1982), and was calculated as

$$\mathrm{d} = 1 - \frac{\sum_{i=1}^{N} (M_i - L_i)^2}{\sum_{i=1}^{N} \left(\left| M_i - \overline{M} \right| + \left| L_i - \overline{M} \right| \right)^2}, \tag{11}$$

where M_i and L_i are respectively the i^{th} manual and LiDAR measurements, with N total measurements (here, $N = 6$), and an overbar denotes an average over all measurements.

The nRMSE and mean bias are respectively calculated as

$$\text{nRMSE} = \frac{\left[\frac{1}{N}\sum_{i=1}^{N}\left(M_i - L_i\right)^2\right]^{1/2}}{\frac{1}{N}\sum_{i=1}^{N} M_i}, \tag{12}$$

and

$$\text{bias} = \frac{1}{N}\sum_{i=1}^{N}\left(L_i - M_i\right). \tag{13}$$

3.2.1. Leaf area distribution Figure 8 shows a comparison between the manual and LiDAR measurements of leaf area for the six zones considered in the experiment. Table 2 presents the data in tabular form, and gives corresponding average parameter values used to calculate a_L. Overall, agreement was very good, with the calculated index of agreement having a value of 98% and nRMSE having a value of 13%. There was a slight positive bias in the LiDAR measurement, with the LiDAR giving a leaf area that was on average 0.3 m^2 larger than that of the manual measurement. Given the available data, we cannot conclude that there is in fact a true bias in the LiDAR measurement, as there are numerous sources of error in the manual measurements. There are small sources of error in defining the tree zones in the field, stripping and transporting the leaves, and in the optical leaf area measurement device. Given the results of Sect. 3.1, one would expect, if anything, a negative bias in the LiDAR measurement. Since the overall bias is quite small, we speculate that the error between the LiDAR and manual measurements is likely within the error bounds for either the LiDAR or manual measurement techniques themselves, although this cannot be conclusively demonstrated given the available data.

Equations 6 (point quadrat theory) and 8 (simplified averaging scheme) were also used to calculate the leaf area from the LiDAR data (Fig. 9, Table 3). Both approaches consistently resulted in lower predictions of leaf area compared to Eq. 10, with the point quadrat approach always giving the lowest value. It is also notable that the difference in the prediction of Eq. 6 and Eq. 10 tends to increase as leaf area density is increased. Both of these results are expected for the reasons discussed above in Sect. 2.5, and based on the results shown in Sect. 3.1.

Given the error metrics listed in Table 3, all methods perform very well and differences in error metrics are minimal. Based purely on the values in Table 3, one may be led to conclude that Eq. 8 is preferable and gives the best performance. Given the theoretical arguments in Sect. 2.5 and the synthetic tests in Sect. 3.1, we find this conclusion to be questionable and suspect that this is partially due to a slight over prediction by the LiDAR measurement (due to the combined error of all measurements involved), which is coincidentally cancelled to some degree by the consistent negative bias introduced by Eqns. 6 and 8.

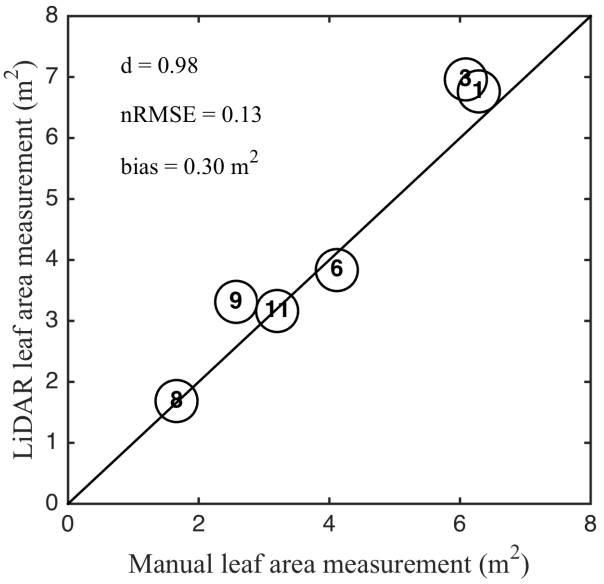

Figure 8. Comparison of the manual leaf area measurements with the LiDAR measurements. Points correspond to the measured leaf area within a voxel and are labeled by the voxel number. The solid diagonal line corresponds to 1:1 or perfect agreement. Agreement is quantified by the index of agreement (d), the normalized root-mean squared error (nRMSE), and mean bias. Note that for the LiDAR measurement leaf area is the product of a_L and the volume of the voxel.

Voxel	$\overline{P}$	$\overline{r}$	$\overline{G}$	LiDAR leaf area (m^2)	manual leaf area (m^2)
1	0.72	1.15	0.20	6.76	6.29
4	0.74	1.05	0.19	6.96	6.09
6	0.90	1.16	0.24	1.68	1.66
7	0.81	1.11	0.23	3.84	4.12
9	0.78	1.02	0.33	3.31	2.57
12	0.79	1.06	0.31	3.16	3.20

Table 2. Parameter values associated with leaf area calculations for experimental data validation study. Note that leaf area can be converted to leaf area density a_L by dividing by the voxel volume (4.4 m^3).

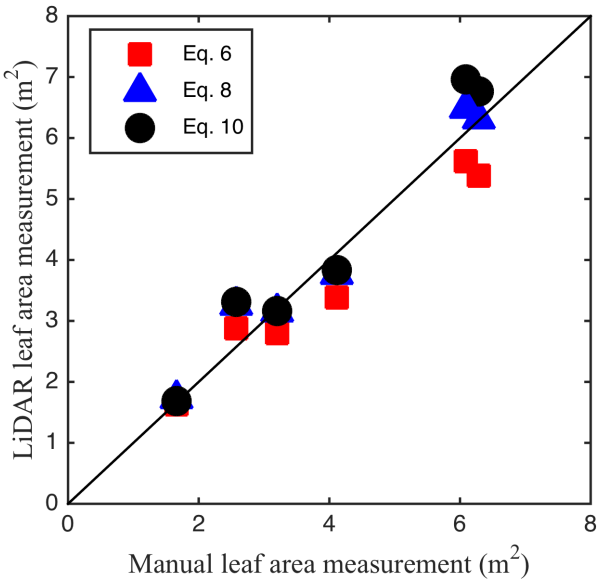

Figure 9. Comparison of the manual leaf area measurements with the LiDAR measurements of leaf area for three different methods of inversion used to calculate leaf area.

Method	d	nRMSE	bias (m^2)
Eq. 6	0.97	0.14	-0.37
Eq. 8	0.99	0.09	0.13
Eq. 10	0.98	0.13	0.30

Table 3. Agreement metrics between three different methods for inversion of leaf area from LiDAR data, which correspond to the points plotted in Fig. 9. Agreement is quantified by the index of agreement (d), the normalized root-mean-squared-error (nRMSE), and mean bias.

4. Summary

A method was developed that uses terrestrial LiDAR scanning data to rapidly and simultaneously measure the leaf area and leaf orientation p.d.f. for an arbitrary volume of leaves. The method uses the triangulation algorithm of Bailey and Mahaffee (2017) to calculate leaf orientations from the scan data, and uses a new technique to directly calculate the fraction of projected leaf area G. The formulation used to calculate leaf area from LiDAR data differs from previous approaches in that it is based on more generalized radiation transfer theory rather than 'point quadrat' theory, which is the method commonly used in previous work (e.g., Radtke and Bolstad, 2001; Hosoi and Omasa, 2006; Béland et al., 2011).

The measurement technique was validated by comparing LiDAR-measured leaf area

against 1) 'synthetic' or computer-generated LiDAR data, and 2) to manually measured leaf area in the field. Agreement between LiDAR and reference measurements was very good. The average normalized root-mean-squared-error for the synthetic data for all densities considered was 15%, and was 13% for the measurements in the field. Leaf area was also calculated from the LiDAR data using 'point quadrat' theory, which consistently predicted lower leaf area compared to when radiation transfer theory was used, an under prediction that continuously increased with increasing leaf area density. For the synthetic data this under prediction was as large as 35% on average for the highest leaf density considered.

Acknowledgments

The authors are indebted to the numerous individuals that assisted in collecting the field data presented in this work: Jesse Mitchell, Tara Neill, Caroline Press, Brenda Shaffer, Lindsey Thiessen, Matt Trappe, Bailey Ward, and Brent Warneke. This research was financially supported by United States Department of Agriculture (USDA) project 5358-22000-039-00D. LiDAR equipment was provided by a grant from the Erath Foundation. The use, trade, firm, or corporation names in this publication are for information and convenience of the reader. Such use does not constitute an endorsement or approval by the USDA or the Agricultural Research Service of any product or service to the exclusion of others that may be suitable.

References

Bailey, B. N. and Mahaffee, W. F. (2017). Rapid measurement of the three-dimensional distribution of leaf orientation and the leaf angle probability density function using terrestrial lidar scanning. *Remote Sens. Env.* In Press. doi:10.1016/j.rse.2017.03.011

Bailey, B. N., Overby, M., Willemsen, P., Pardyjak, E. R., Mahaffee, W. F., and Stoll, R. (2014). A scalable plant-resolving radiative transfer model based on optimized GPU ray tracing. *Agric. For. Meteorol.*, 198-199:192–208.

Bailey, B. N. and Stoll, R. (2016). The creation and evolution of coherent structures in plant canopy flows and their role in turbulent transport. *J. Fluid Mech.*, 789:425–460.

Bailey, B. N., Stoll, R., Pardyjak, E. R., and Miller, N. E. (2016). A new three-dimensional energy balance model for complex plant canopy geometries: Model development and improved validation strategies. *Agric. For. Meteorol.*, 218-219:146–160.

Béland, M., Baldocchi, D. D., Widlowski, J.-L., Fournier, R. A., and Verstraete, M. M. (2014). On seeing the wood from the leaves and the role of voxel size in determining leaf area distribution of forests with terrestrial lidar. *Agric. For. Meteorol.*, 184:82–97.

Béland, M., Widlowski, J.-L., Fournier, R. A., Côté, J.-F., and Verstraete, M. M. (2011). Estimating leaf area distribution in savanna trees from terrestrial LiDAR. *Agric. For. Meteorol.*, 151:1252–1266.

Bréda, N. J. J. (2003). Ground-based measurements of leaf area index: a review of methods, instruments and current controversies. *J. Exp. Botany*, 54:2403–2417.

Caldwell, M. M., Harris, G. W., and Dzurec, R. S. (1983). A fiber optic point quadrat system for improved accuracy in vegetation sampling. *Oecologia*, 59:417–418.

Fong, W.-K., Matsumoto, H., and Lun, Y.-F. (2009). Application of system dynamics model as decision making tool in urban planning process toward stabilizing carbon dioxide emissions from cities. *Building and Env.*, 44:1528–1537.

Henning, J. G. and Radtke, P. J. (2006). Ground-based laser imaging for assessing three-dimensional forest canopy structure. *Photogramm. Eng. Remote Sens.*, 72:1349–1358.

Hosoi, F. and Omasa, K. (2006). Voxel-based 3-D modeling of individual trees for estimating leaf area density using high-resolution portable scanning lidar. *IEEE Trans. Geosci. Remote Sens.*, 40:3610–3618.

Jonckheere, I., Fleck, S., Nackaerts, K., Muys, B., Copping, P., Weiss, M., and Baret, F. (2004). Review of methods for in situ leaf area index determination Part I. Theories, sensors and hemispherical photography. *Agric. For. Meteorol.*, 38:33–44.

Jones, H. G. and Vaughan, R. A. (2010). *Remote Sensing of Vegetation: Principles, Techniques, and Applications*. Oxford University Press, Oxford, UK. 400 pp.

Kerr, R. A. (2012). Weather forecasts slowly clearing up. *Science*, 338:734–737.

Kobayashi, H., Baldocchi, D. D., Ryu, Y., Chen, Q., Ma, S., Osuna, J. L., and Ustin, S. L. (2012). Modeling energy and carbon fluxes in a heterogeneous oak woodland: A three-dimensional approach. *Agric. For. Meteorol.*, 152:83–100.

Krayenhoff, E. S., Christen, A., Martilli, A., and Oke, T. R. (2014). A multi-layer radiation model for urban neighbourhoods with trees. *Boundary-Layer Meteorol.*, 151:139–178.

Lemonsu, A., Masson, V., Shasua-Bar, L., Erell, E., and Pearlmutter, D. (2012). Inclusion of vegetation in the town energy balance model for modelling urban green areas. *Geosci. Model Dev.*, 5:1377–1393.

Lindberg, F. and Grimmond, C. S. B. (2011). The influence of vegetation and building morphology on shadow patterns and mean radiant temperatures in urban areas: model development and evaluation. *Theor. Appl. Cim.*, 105:311–323.

Modest, M. F. (2013). *Radiative Heat Transfer*. Academic Press, Waltham, MA, third edition. 904 pp.

Norman, J. M. and Campbell, G. S. (1989). Canopy structure. In Pearcy, R. W., Ehleringer, J., Mooney, H. A., and Rundel, P. W., editors, *Plant Physiological Ecology: Field Methods and Instrumentation*, pages 301–326. Chapman & Hall, New York.

Olchev, A., Radler, K., Sogachev, A., Panferov, O., and Gravenhorst, G. (2009). Aplication of a three-dimensional model for assessing effects of small clear-cuttings on radiation and soil temperature. *Ecol. Model.*, 220:3046–3056.

Pisek, J., Sonnentag, O., Richardson, A. D., and ottus, M. M. (2013). Is the spherical leaf inclination angle distribution a valid assumption for temperate and boreal broadleaf tree species? *Agric. For. Meteorol.*, 169:186–194.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (2007). *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, Cambridge, U.K. 1256 pp.

Radtke, P. J. and Bolstad, P. V. (2001). Laser point-quadrat sampling for estimating foliage-height profiles in broad-leaved forests. *Can. J. For. Res.*, 31:410–418.

Rosell, J. R., Llorens, J., Sanz, R., Arnó, J., R.-D., M., Masip, J., Escolà, A., Camp, F., Solanelles, F., Gràcia, F., Gil, E., Val, L., Planas, S., and Palacín, J. (2009). Obtaining the three-dimensional structure of tree orchards from remote 2D terrestrial LIDAR scanning. *Agric. For. Meteorol.*, 149:1505–1515.

Ross, J. (1981). *The Radiation Regime and Architecture of Plant Stands*. Dr. W. Junk Publishers, The Hague, The Netherlands. 424 pp.

Searchinger, T., Heimlich, R., Houghton, R. A., Dong, F., Elobeid, A., Fabiosa, J., Tokgoz, S., Hayes, D., and Yu, T.-H. (2008). Use of u.s. croplands for biofuels increases greenhouse gases through emissions from land-use change. *Science*, 319:1238–1240.

Sinoquet, H., Le Roux, X., Adam, B., Ameglio, T., and Daudet, F. A. (2001). RATP: A model for simulating the spatial distribution of radiation absorption, transpiration and photosynthesis within canopies: application to an isolated tree crown. *Plant Cell Environ.*, 24:395–406.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D., Barker, D., Wang, W., and Powers, J. G. (2005). A description of the Advanced Research WRF Version 2. Technical Report NCAR/TN-468+STR, NCAR.

Suffern, K. G. (2007). *Ray Tracing from the Ground Up*. A K Peters/CRC Press, Boca Raton, FL. 784 pp.

Warren Wilson, J. (1960). Inclined point quadrats. *New Phytol.*, 59:1–8.

Welles, J. M. and Norman, J. M. (1991). Instrument for indirect measurement of canopy structure. *Agron. J.*, 83:818–825.

Willmott, C. J. (1981). On the validation of models. *Phys. Geogr.*, 2:184–194.

Willmott, C. J. (1982). Some comments on the evaluation of model performance. *Bull. Amer. Meteor. Soc.*, 63:1309–1313.

Wu, H., He, Z., and Gong, J. (2010). A virtual globe-based 3D visualization and interactive framework for public participation in urban planning processes. *Comp., Env. Urban Sys.*, 34:291–298.

Yang, X., Strahler, A. H., Schaaf, C. B., Jupp, D. L. B., Yao, T., Zhao, F., Wang, Z., Culvenor, D. S., Newnham, G. J., Lovell, J. L., Dubayah, R. O., Woodcock, C. E., and Ni-Meister, W. (2013). Three-dimensional forest reconstruction and structural parameter retrievals using a terrestrial full-waveform lidar instrument (Echidna®). *Remote Sens. Env.*, 135:36–51.